

Habitat-Related Variability in the Cave-Dwelling
Minnow, *Hybopsis harperi*

By
JOHN F. HOWELL

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
June, 1960

ACKNOWLEDGEMENTS

No complete list of acknowledgements is possible for this report; it would have to begin with a mother who loves nature and extend to a wife who patiently copies tables and reads proof. In between would be found participants in bull sessions, English teachers, and a number of just plain friends. In addition to the people mentioned below, I extend my thanks to all of these.

I would like to specifically thank Dr. Archie F. Carr, who stimulated my interest in Hybopsis and who has served as chairman of my advisory committee, and the other members of the committee, both past and present: Dr. D. B. Duncan, now at the University of North Carolina; Mr. Victor Chew, now at North Carolina State College; Dr. A. B. Grobman, now Director of the A. I. B. S. Curriculum Study Committee; Dr. A. T. Wallace; Dr. E. R. Jones; Dr. H. M. Wallbrunn; and Dr. John D. Kilby, whose assistance has gone beyond that expected of a committee member.

The computations involved in this work would have been impossible except for the assistance of the Statistical Laboratory of the University of Florida. I especially want to thank Mr. Carlis Taylor, Mr. Edward Koontz, and Mr. H. R. Dean for their assistance in pro-

gramming and using the computer. Mr. B. W. Brentlinger, of the IBM Supplies Division Card Plant in Sherman, Texas, also rendered invaluable aid in my use of the IBM accounting machine at that plant.

For financial assistance during the time that I have worked on this problem, I am indebted to the National Science Foundation, Austin College, the Graduate Council and College of Arts and Sciences of the University of Florida, and to my wife Gwen W. Howell.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF TABLES	v
LIST OF FIGURES	vii
INTRODUCTION	1
MATERIALS AND METHODS	
Samples	7
Counts and Measurements	11
Computations	14
SINGLE FACTOR ANALYSIS	
Analysis of Variance	15
Meristic characters	19
Angle of gape	20
Measurements	21
Graphic Analysis of angle of gape	27
MULTIVARIATE ANALYSIS	
Mahalanobis' Generalized Distance Function	29
Geometric Derivation of D^2	31
Computation of D^2	34
Interpretation of D^2	34
DISCUSSION AND CONCLUSIONS	42
SUMMARY	50
LITERATURE CITED	53
APPENDIX	57
BIOGRAPHICAL SKETCH	84

LIST OF TABLES

Table		Page
1	Interpretation of the orthogonal comparisons	17
2	Variance ratio (<u>F</u>) values for comparisons of meristic characters	19
3	Variance ratio (<u>F</u>) values for comparisons of measurements	22
4	Values of <u>D</u> ² , corrected for bias	35
5	Intracluster and intercluster values for <u>D</u> ² in the clustering configuration "A"	37
6	Intracluster and intercluster values for <u>D</u> ² in the clustering configuration "B"	40
7	Means and standard deviations for all characters in all samples	59
8	Values of Chi Square for Bartlett's test	62
9	Coefficients and divisors for computing sums of squares for orthogonal comparisons	63
10	Variance-covariance matrix for measurements and ratios	64
11	Inverse of variance-covariance matrix	65
12	Analysis of variance for dorsal rays	66
13	Analysis of variance for anal rays	67
14	Analysis of variance for scales in lateral line	68
15	Analysis of variance for scales above lateral line	69

LIST OF TABLES--Continued

Table		Page
16	Analysis of variance for scales below lateral line	70
17	Analysis of variance for angle of gape	71
18	Analysis of variance for length	72
19	Analysis of variance for depth of body	73
20	Analysis of variance for head length	74
21	Analysis of variance for snout length	75
22	Analysis of variance for width of eye	76
23	Analysis of variance for length to dorsal fin	77
24	Analysis of variance for depth of caudal peduncle	78
25	Analysis of variance for width of head	79
26	Analysis of variance for head length / snout length	80
27	Analysis of variance for head length / eye width	81
28	Analysis of variance for head length / head width	82
29	Analysis of variance for peduncle depth / depth of body	83

LIST OF FIGURES

Figure	Page
1 Map of collection sites in North Central Florida	8
2 Population-range diagram for angle of gape	28
3 Cluster configurations "A" and "B" with average intercluster values for <u>D</u>	39

INTRODUCTION

Hybopsis harperi is a rheotropic cyprinid minnow found in Florida in two extremely different situations and in several variants of each. Although spring-fed runs and rivers constitute its main habitat, H. harperi is also found in small, steep-sided sink holes in the Ocala limestone of North Central Florida. Hellier (1957), in his study of the Santa Fe River, pointed out that it is the most abundant fish in the spring-fed stretches of the river. Marshall (1941, 1947) discussed the seemingly strong ecologic divergence of the two habitats occupied.

Marshall considered the absence of Gambusia from most sink holes as evidence for dispersal of Hybopsis through a subterranean water system. The presence of a large population of H. harperi in a pond in a lime rock mine near Haile, Florida, further suggests this means of dispersal. Within a year of its filling with water, this pond contained Hybopsis but no Gambusia, even though it was undoubtedly suited for occupation by Gambusia. The ecological suitability of the habitat is implied by the fact that a local bass fisherman had succeeded in introducing Xiphophorus variatus into the pond in order to have a ready suppl' of bait. Xiphophorus, a poeciliid not native to Florida, appeared to be occupying the same niche that Gambusia does in sink

holes into which it has been introduced.

If dispersal of H. harperi occurs through the underground system, there is certainly some degree of isolation among the subpopulations in the sink holes, and also between these and the ones in streams. In fact, it is conceivable that in some cases there may be a greater degree of isolation between neighboring sink holes than between these and some distant stream.

Therefore, unless we are to postulate some directed migration among the various populations, H. harperi in northern Florida is evidently a large population composed of semi-isolated subpopulations. With this population structure we would expect, in the presence of selection pressure, to find adaptive changes in the cave and sink hole populations. Moreover, under Wright's (1931) hypothesis, these changes should occur rather rapidly. In the absence of selection pressure, random changes in gene frequencies would be expected to manifest themselves in alterations of morphometric or meristic characters. Also, with the degree of homozygosity which results from inbreeding in isolated subpopulations, there might be an increase in nongenetic variability (Lerner, 1954).

Thus, there are two factors which could be expected, a priori, to contribute to the formation of subspecifically distinct populations. First, one can offer H. harperi as a model for the situation in which

Wright (1931) suggested that evolution would proceed most rapidly; i.e., a large population broken up into small, almost (but not completely) isolated units. Secondly, if there is no migration at all between the subpopulations, one would expect genetic drift to have allowed some nonadaptive differences to develop.

Under the first of these hypotheses, one can even postulate a direction for this presumptively rapid evolution. Sink holes such as those included in this study form a graded series from fairly well-lighted ones to completely dark caves. Thus the population structure of Hybopsis could assist in the bridging of the gap between the "adaptive peak" corresponding to the stream habitat and that corresponding to the cave or sink hole habitat. No blind or unpigmented Hybopsis, which might be expected if evolution had proceeded far enough in the direction suggested, have been found.

In addition to the nonadaptive divergence of populations caused by random genetic drift, "accidents of sampling" might also be expected in the colonization of sink holes, or in the re-establishing of populations from survivors of periodic reductions during times of drought.

Regardless of the causes of variation in Hybopsis harperi, there would seem to be grounds to expect that a study of this variation would reveal differences sufficient to justify separation of at least

two subspecies. Hubbs and Crowe (1956) did this, describing the sink hole form as a distinct subspecies, Hybopsis harperi subterranea. This fact and the pattern of local distribution of Hybopsis in northern Florida suggest that a further study of variation in this species might shed light on the effect of population structure on subspeciation and on the course of microevolution generally.

Conventional practices with respect to the naming of subspecies have been subjected to intensive examination in recent years. Beginning with a paper by Wilson and Brown (1953), a series of somewhat polemic essays has appeared in Systematic Zoology, with such proposals as that trinomens be stripped of both priority and italics (Gosline, 1954) or that all subspecies be completely abolished (Brown and Wilson, 1955). There has even been a recommendation that an "objective" subspecies definition, based solely on geographic considerations, be adopted (Edwards, 1956), but Borgmeier (1957) favored a morphologically conceived subspecies. Hubbell (1954), summarizing a symposium on the subject, concluded that in general infraspecific variation is best treated by description, graphic representation, and nontechnical names, and not by formal names subject to the rules of nomenclature.

Except for some purely legalistic questions of nomenclature, criticism of the current subspecies concept seems to center around three

main issues, as follows: (1) the criteria for distinguishing subspecies are usually quantitative characters which do not exhibit complete discontinuity; (2) analysis of several characters frequently shows that there is nonconcordance in the variation of these characters; and (3) the adoption of a "percent rule," which is made necessary by the lack of discontinuity, is altogether arbitrary.

Two recent publications, neither of them in the mainstream of the subspecies controversy, have dealt with the statistical aspects of the problem. Pimentel (1958) suggested the study of many characters and material from many localities before reaching a conclusion concerning the probable existence of subspecies. He also outlined a sound statistical procedure for such a study. Marr (1957) defined several levels of population units in fishes and proposed methods for studying the subpopulations.

The present study stands somewhere between these two papers in its approach to the problem of subspeciation in Hybopsis. Pimentel, who was concerned primarily with morphological data, was especially interested in the assured randomness of samples. Although he encouraged inclusion of many characters, his methods were restricted to analysis of each character separately. Marr, on the other hand, extended his techniques of analysis to include biochemical and tagging data and genetic studies wherever they are possible. He also

suggested the use of multivariate analysis in the morphometric phase of an investigation, and Royce (1957), in the same symposium, developed this technique.

My study has utilized some specimens from museum collections and others collected in much the same way as those that furnish the material of the usual ichthyological investigation. Therefore, in contrast to Pimentel's suggestion, randomness of samples is not assured. Even though the populations of variables are thus non-randomly sampled, there is no reason to believe that the departures from randomness are the same for all variables; so the synthesis of a composite multivariate statistic, such as Mahalanobis' D^2 , should be less affected by the nonrandomness than single variable statistics.

My method of evaluating variation in subpopulations of Hybopsis harperi has been to use the analysis of variance technique, as suggested by Pimentel, on single characters and then to extend the analysis to multivariate methods, as suggested by Marr and by Royce.

MATERIALS AND METHODS

Samples

The data compiled for this study consist of counts and measurements from a total of 635 specimens of Hybopsis harperi, all collected in North Central Florida (Figure 1). Some of the specimens were in the University of Florida-Florida State Museum collections; the rest were collected by the author. For the purpose of analysis it will be convenient to recognize three collection-lots in the material: a Special Series, a Santa Fe Series, and a Main Series. In the following gazetteer of localities for these collections, the abbreviation by which each will hereinafter be designated is given first. The first two letters of the abbreviation indicate the location, as shown in Figure 1; the third letter indicates whether the habitat is standing water or running water, and a number designates one of several samples collected at the same place.

The Special Series comprises a random sample of 100 specimens from each of the following three collections:

SFR-1: A collection (UF-8425) made by Hellier in the
Santa Fe River near Duncan's Landing on
November 17, 1956.

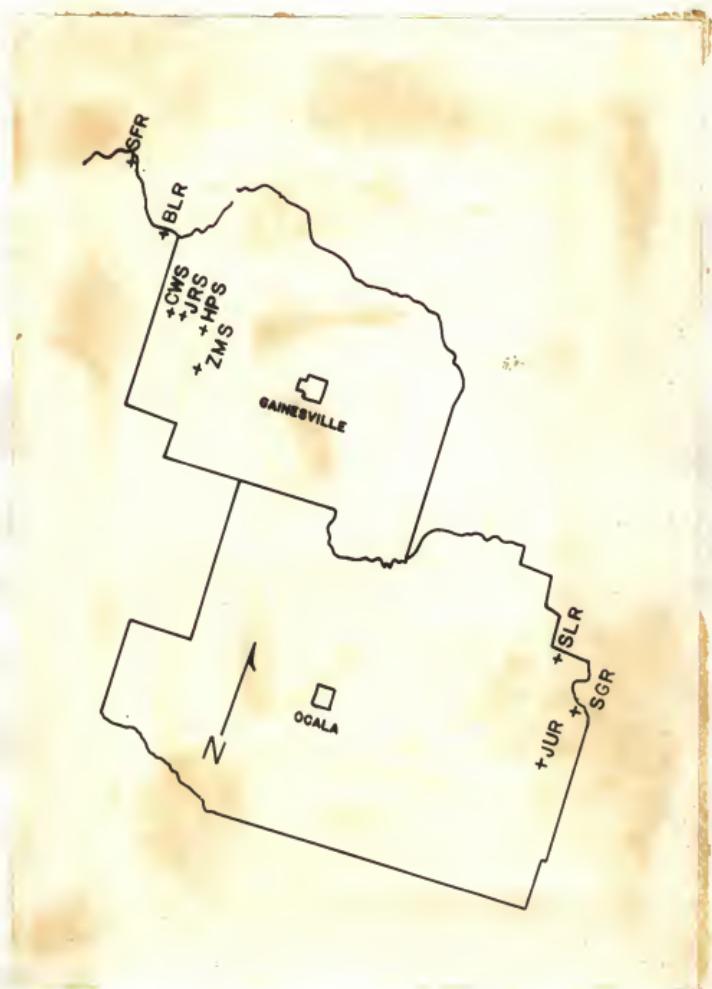


Figure 1. Map of collection sites in North Central Florida.

HPS: A collection made from a lime rock mine near Haile on January 29, 1959. This pool became populated by Hybopsis within a year after the mining operation reached the water table, and the environment is quite unlike the stream habitat or the sink hole habitat.

ZMS: A collection from Zamia Sink, a sink hole in Alachua County, made on December 29, 1958.

The Santa Fe Series is made up of five more samples from the same section of the Santa Fe River as SFR-1. These collections, also made by Hellier during the course of his study of the Santa Fe, are as follows:

SFR-2: A collection made February 25, 1956, by means of a common sense seine and a ten-foot bag seine.

SFR-3: A collection made July 3, 1955, with a common sense seine and a ten-foot bag seine.

SFR-4: A collection made July 7, 1956, with a fifty-foot bag seine and a common sense seine.

SFR-5: A collection made on July 22, 1956, with a fifty-foot bag seine.

SFR-6: A collection made September 14, 1956, using a fifteen-foot bag seine.

Collection SFR-3 contained only 25 specimens, all of which were included in the sample. Each of the others was a random sample of 30 specimens from one of the above collections.

The Main Series is made up of samples from a variety of stream and sink hole environments in North Central Florida. Each is a random sample of 30 specimens from one of the collections indicated below:

JRS-1: A collection made on July 19, 1959, from Jerome Sink, a sink hole in Alachua County. This is the type locality of Hybopsis harperi subterranea Hubbs and Crowe.

JRS-2: A collection from Jerome Sink which was made May 30, 1955.

CWS: A collection from Cow Sink which was made January 29, 1959. Cow Sink is located about 200 yards west of Jerome Sink.

BLR: A collection from Blue Springs in Gilchrist County on June 11, 1959. Blue Springs is a tributary of the Santa Fe River.

SGR: A collection made from Silver Glen Springs on September 20, 1951 (UF-6307).

SLR: A collection from Salt Springs on November 9, 1949
(UF-6303).

JUR: A collection from Juniper Springs made by Dr.
John D. Kilby on November 1, 1958.

Counts and Measurements

The values for five meristic characters were determined for the specimens in the Special Series and in all samples of the Main Series except SGR and SLR. These characters and their abbreviations are as follows: dorsal fin rays (D), anal fin rays (A), scales in the lateral line (LL), scales above the lateral line (SA), scales below the lateral line (SB). An analysis of variance for the meristic characters is presented below. Counts were discontinued in the later stages of the study after a preliminary analysis of the Special Series showed no reliable differences in the meristic characters in these three samples. Values for thirteen morphometric characters were obtained for each specimen. The meristic counts were taken according to Hubbs and Lagler (1941), and wherever possible the measurements were taken between the reference points suggested by these authors.

The morphometric characters are as follows:

AG-----Angle of gape

L----- Standard length
DB----- Depth of body
HL----- Head length
SL-----Length of snout
E-----Width of eye
LD----- Length to dorsal fin
PD----- Depth of caudal peduncle
HW----- Width of head
HL/SL--Length of head/length of snout
HL/E---Length of head/width of eye
HL/HW- Head length/head width
PD/DB--Depth of peduncle/depth of body.

All the linear measurements except L, DB, and LD were made with the aid of a camera lucida. The image of the fish was projected onto a sheet of white paper on which index lines had been previously inscribed, and the reference points were marked on the index lines. The measurements were then taken from the sheet with dividers. The magnification of the projected image was 14.2 diameters. This method was not applicable for L, DB, or LD because the reference points for these measurements could not be included in one field. An opaque projector, magnifying 1.57 diameters, was used to project

these reference points onto another index line on the same sheet. These points were also marked and the measurements taken with dividers and the same scale.

For the purpose of this statistical study, the actual values of the measurements were not determined; the computations were performed on measurements in fiftieths of an inch, taken directly from the data sheets. To convert these to actual metric values in millimeters, one may multiply L, DB, and LD by 0.3236, or the other measurements by 0.0358. Obviously the statistical results would not change with this transformation. In order to facilitate comparisons, Table 7 expresses the means and standard deviations of all samples in millimeters.

The reference points for the angle of gape (AG) were also marked by projection through the camera lucida. The angle measured was that formed by the "horizontal" index line and a line drawn between the anteriormost and posteriormost extent of the gape. The difficulty of making a true horizontal superimposition of the index line produced rather large variations in this measurement. The values for AG reported here probably can not be compared with those reported by Hubbs and Crowe (1956) because of differences in the method of measuring the angle.

Computations

Computations of the sums of squares, sums of products, means, and standard deviations for each sample were performed on the IBM 650 computer at the University of Florida Statistical Laboratory. These were summarized by use of the 407 accounting machine at the IBM Supplies Division Card Plant at Sherman, Texas. Inversion of the variance-covariance matrix was done by the Statistical Laboratory, as were the matrix multiplications necessary for the computation of D^2 .

SINGLE FACTOR ANALYSIS

Analysis of Variance

Since multivariate statistical analysis requires computations that are laborious even on modern electric desk calculators, single factor analyses should be used if they will allow valid conclusions to be drawn. Therefore, the analysis of variance technique was used here in an attempt to identify any of the characters which might allow discrimination among the subpopulations on the basis of single factors. In addition to the usual analysis of variation between samples generally, I have also been able to make orthogonal comparisons in a search for the source of variations between samples. The one odd-sized sample in the Santa Fe Series prevents complete orthogonality of any meaningful comparisons. Thus, the fourth comparison is merely a measure of the variation remaining after the other comparisons have been made. That a standard sample size should be adopted in variational investigations is indicated by the fact that much of the significant variation was in that remainder.

Although meristic, scalar, linear, and angular characters may be included in the same multivariate analysis (Olsen and Miller, 1958), preliminary analyses of the Special Series and the differences in units

led to a decision to restrict the analysis of the meristic characters to the single factor analysis of variance, including orthogonal subdivision of the between-sample sums of squares.

One of the assumptions in the analysis of variance is that the variances in the different samples are homogeneous. Departures from this assumption will result in a loss of sensitivity in the analysis (Cochran, 1947). Even though this technique was to be used primarily for identifying characters showing variation, Bartlett's test of the homogeneity of variances (Ostle, 1954) was performed. In all three series, the hypothesis of homogeneity of the variances was rejected for most of the characters (Table 8, Appendix). Li (1957), Snedecor (1956), and Cochran and Cox (1957) suggest that moderate departures from homogeneity are not serious in analysis of biological material, but they all caution the worker to be careful in interpreting the significance level of the differences. In this connection, it is interesting to note that, of 161 F values which exceeded the 5 percent tabular value in these variance analyses, 130 also exceeded the 0.1 percent level. One may then cautiously suggest that these F values indicate real differences in the samples on which they are based. Throughout this report the values which exceed the 5 percent, 1 percent, and 0.1 percent values for F will be indicated respectively by single, double, and triple asterisks.

TABLE I

INTERPRETATION OF THE ORTHOGONAL COMPARISONS

Comparison	Series	A significant F value indicates a difference between the:
C ₁	Special	sample from running water and the two samples from still water.
C ₂	Special	two samples from standing water.
C ₃	Santa Fe	samples collected with a fifty-foot bag seine and those taken with shorter seines.
C ₄	Santa Fe	two samples collected in the same area during the same month.
C ₅	Santa Fe	samples collected in February and the one taken in September.
C ₆	Santa Fe	samples in this series which is not accounted for by the other three comparisons.
C ₇	Main	samples from springs and those from sink holes.
C ₈	Main	samples from two closely adjacent sink holes, Jerome Sink and Cow Sink.
C ₉	Main	two samples taken from Jerome Sink at different times.
C ₁₀	Main	sample from Blue Springs and the three from Marion County springs. (For the meristic characters, this comparison is only between Blue Springs and Juniper Springs.)

TABLE 1-Continued

Comparison	Series	A significant <u>F</u> value indicates a difference between the:
C ₁₁	Main	sample from Juniper Springs and the other two from Marion County.
C ₁₂	Main	sample from Salt Springs and that from Silver Glen Springs.

In the Special Series, two single-degree-of-freedom orthogonal comparisons were possible. Sums of squares for these and all other comparisons were computed using the method of coefficients (Ostle, 1954; p. 267). The first Special Series comparison (indicated by C₁) tests the differences in the characters which are related to the differences in standing- or running-water habitats. The second comparison (C₂) compares the two standing-water samples, from Zamia Sink and the lime rock pit at Haile. A similar, but more extensive, analysis could be made for the Main Series. In this series, C₇ measures differences between springs and sink holes; C₈ compares two sink holes; C₉ compares samples from the same sink hole in different years; and C₁₀ compares two springs. The notation for all orthogonal comparisons is consistent throughout this report, and the coefficients for these comparisons are presented in the Appendix, Table 9. Table 1 presents the whole list of comparisons and their interpretation.

Meristic characters. Table 2 summarizes the results of the analyses of meristic characters. It may readily be seen that, while differences related to the different environments are indicated for SA and possibly LL, there are also differences in LL in samples from closely adjacent sink holes and even in two samples from the same sink hole. Thus it seems that only SA supports the separation of the two named subspecies.

TABLE 2

VARIANCE RATIO (F) VALUES FOR COMPARISONS OF MERISTIC CHARACTERS

Char- acter	Special Series			Main Series		
	C ₁	C ₂	C ₇	C ₈	C ₉	C ₁₀
D	1.38	0.17	0.00	0.51	0.17	0.00
A	1.87	16.76***	0.30	0.08	0.23	0.93
LL	8.85***	33.08***	1.04	12.62***	24.91***	1.35
SA	146.94***	11.85***	15.51***	0.00	1.58	0.40
SB	0.06	6.56*	0.12	0.04	2.83	0.45

Counts of meristic characters were discontinued after analysis of the Special Series because the results indicated that these data were not of value in distinguishing the subspecies. During the course of my counting of scales above the lateral line, I observed that a very slight longitudinal displacement of the dorsal fin would result in the increase

or decrease of the scale count by one scale. Thus the variation observed between samples from the standing-water and running-water habitats might well have been due to variation in the location of the dorsal fin. This character (LD) does show significant variation between the two habitats, but an analysis of covariance of the scale count on LD did not reduce the value of the variance ratio to a nonsignificant level (Special Series, $F = 58.10^{***}$; Main Series, $F = 4.56^*$).

Angle of gape. Since the meristic characters were not to be included in the multivariate analysis, and since the angle of gape (AG) was the only other character for which the units were not either linear or ratios of linear units, this character was also analyzed separately. The analysis of the Special Series showed that there was no significant variation between the samples collected in still water and the one collected in running water ($F = 2.38$) and that the variance ratio for the comparison of the two still-water samples only exceeded the tabulated F value for a 5 percent significance level.

For the Santa Fe Series, one orthogonal comparison (C_3) measures variation between material collected with a fifty-foot bag seine and that taken with short seines. Another comparison (C_4) compares two different collections made in the same month, and a third (C_5) compares a collection made in February with one made in September.

The fourth orthogonal component of the variation (the comparison C_6) was chosen as the remaining variation because the unequal sample sizes precluded any other meaningful comparison. The results of this analysis, which are presented in Table 18, indicated no significant differences in any of these components.

In the Main Series, six meaningful comparisons are possible: C_7 , C_8 , and C_9 are the same as they were for the meristic characters. With the added data of SGR and SLR, C_{10} now compares the sample from Blue Springs with all three samples from the Marion County springs, instead of with Juniper Springs only. The two additional degrees of freedom allow comparisons among the samples from springs in Marion County: C_{11} compares the sample from Juniper Springs with the other two, and C_{12} compares the samples from Salt Springs and Silver Glen Springs. In this analysis, also presented in Table 18, it appears that there may be some difference in AG in fish from the two habitats and also some geographic difference. However, the variation found in two samples from the same sink hole and the results of the analysis of the Special Series reduce the confidence that can be placed in this character. In addition, as mentioned above, the precision of the measurement involved is open to some question.

Measurements. The comparison C_1 for all measurements in the

TABLE 3
VARIANCE RATIO (*F*) VALUES FOR COMPARISONS OF MEASUREMENTS

Compar- ison	Characters					LD
	L	DB	HL	SL	E	
C ₁	827.66***	12.91***	21.14***	24.31***	140.08***	80.39***
C ₂	1.09	49.52***	0.16	0.99	6.58*	0.03
C ₃	40.14***	77.99***	38.15***	33.76***	19.80***	45.74***
C ₄	2.39	7.29***	0.01	0.002	0.69	1.40
C ₅	0.31	0.67	1.84	0.84	4.75*	0.0004
C ₆	22.91***	12.40***	39.48***	26.38***	21.28***	13.10***
C ₇	11.96***	4.68*	4.90*	2.53	0.89	8.07**
C ₈	234.69***	11.04***	60.02***	62.07***	62.81***	22.86***
C ₉	405.72***	65.66***	45.98***	30.90***	53.62***	55.92***
C ₁₀	39.87***	6.41*	8.67**	7.50**	11.83***	2.34
C ₁₁	311.68***	48.56***	34.05***	59.10***	64.69***	38.18***
C ₁₂	220.09***	17.56***	26.02***	18.83***	42.65***	18.79***

TABLE 3--Continued

Comparisons	Characters					
	PD	HW	HL/SL	HL/E	HL/HW	PD/DB
C1	53.76***	13.34***	12.94***	74.50***	4.45*	13.87***
C2	16.58***	7.42**	6.93**	32.31***	38.19***	48.01***
C3	35.45***	45.24***	4.64*	9.26**	2.97	41.38***
C4	0.84	0.89	0.14	0.80	7.62**	15.74***
C5	6.79*	0.22	16.05***	11.77***	5.85*	31.87***
C6	6.41*	22.81***	0.28	9.01**	19.64***	4.04*
C7	0.52	15.74***	0.36	32.27***	13.32***	13.63***
C8	38.20***	54.85***	17.65***	15.88***	0.003	40.22***
C9	51.03***	48.21***	1.05	10.31**	1.61	13.35***
C10	5.89*	2.37	0.53	1.60	17.81***	0.96
C11	39.22***	29.04***	56.11***	0.97	0.88	18.68***
C12	20.05***	24.09***	4.63*	2.69	1.77	1.33

* P ≤ 0.05
** P ≤ 0.01
*** P ≤ 0.001

Special Series indicates that there are differences between fish from running water and those from still water. For all but one of these measurements, the computed F exceeds the 0.1 percent value in the F-table. Comparison C₂, which compares the lime rock pit sample with that from Zamia Sink, indicates significant differences in eight of the twelve measurements; those which do not exhibit a significant F value are L, HL, SL, and LD.

The complete array of comparisons in the Main Series (Table 3) suggests that there is much local variation in these characters. In Jerome Sink and Cow Sink, which are located about 200 yards apart, there are differences in every character except HL/HW. And in the two samples from Jerome Sink, there are differences in all except that character and HL/SL. On the other hand, the comparison of all spring samples with all the sink hole samples (C₇) shows a difference in HL/HW, but none in HL/SL, SL, E, or PD. Some geographic variation is indicated by the significant F values in C₁₀ for the following measurements: L, HL, SL, E, DB, PD, and HL/HW. For the other two comparisons in this series, Juniper Springs fish were compared with those from Silver Glen Springs and Salt Springs, and the latter two were compared with each other. Between Juniper and the other two springs, there were differences in all the characters except for two ratios, HL/E and HL/HW. The Salt Springs and Silver

Glen Springs samples exhibited differences in all characters but these two and the ratio PD/DB.

From the results of this analysis, one can conclude that, with the possible exception of the ratio of head-length to head-width, none of the measurements taken will allow consistent identification of fish from the two habitats. Even this ratio seems to exhibit some geographic variation.

The results presented above indicate that there is, in samples of Hybopsis, some variation that is related neither to geographic factors nor to differences in the habitat. Since large collections taken from the Santa Fe River, over a period of time and by various methods, were available to the collections of the Florida State Museum, some of these samples were included as the Santa Fe Series in this study. In this analysis an effort was made to relate differences in these measurements to differences in methods or times of collecting.

With the five samples of the Santa Fe Series, four orthogonal components of the variation can be distinguished. The first comparison includes the sample (SFR-3) of 25 individuals taken with short seines because this series was studied for the particular purpose of examining differences caused by variable sampling. Inclusion of this sample precludes orthogonality of a comparison between samples taken at the

same time in different years. The fourth comparison (C_6) is thus a measure of the variation remaining after removal of the variation due to C_3 , C_4 , and C_5 .

The analysis of C_3 showed that there are significant differences in samples taken with different equipment in all of the characters except HL/HW. In the comparison of the two July samples (C_4), only HL/HW, DB, and PD/DB show significant differences. Between the February and September samples there are significant differences in the four ratios and in PD and E as well. Even after the removal of variation due to these three comparisons, there is still significant variation in all of the characters except HL/SL. Taking all measurements of this series collectively, it is evident that there is no measurement included in which there is not some significant variation among this set of five samples.

On the basis of the analysis of variance of the three series of samples, we may conclude that there is no character among the measurements made that will allow consistent separation of Hybopsis in its two habitats. Before proceeding to the multivariate analysis of the twelve characters, it might be instructive to select, for a graphic analysis such as that suggested by Pimentel (1958), that character which most nearly allows diagnosis of the forms in the two habitats.

Graphic Analysis of Angle of Gape

Although, as was pointed out earlier, the precision of the measurement of the angle of gape is open to question, and although it varies considerably within one sink hole, this nevertheless is the best single character for this purpose. Moreover, it was the only character for which Bartlett's test indicated that the variance was homogeneous in both the Main Series and the Special Series, and homogeneity of variance is a prerequisite for Pimentel's method. Figure 2 is a graphic representation of variation in AG for all samples in the Main Series and the Special Series.

It is obvious in this chart that this character will not support recognition of the two subspecies that have been proposed; nonoverlap of the entire open rectangle is required to indicate subspecific differentiation under the interpretation of the 75 percent rule suggested by Pimentel (84 percent of one population different from 84 percent of another).

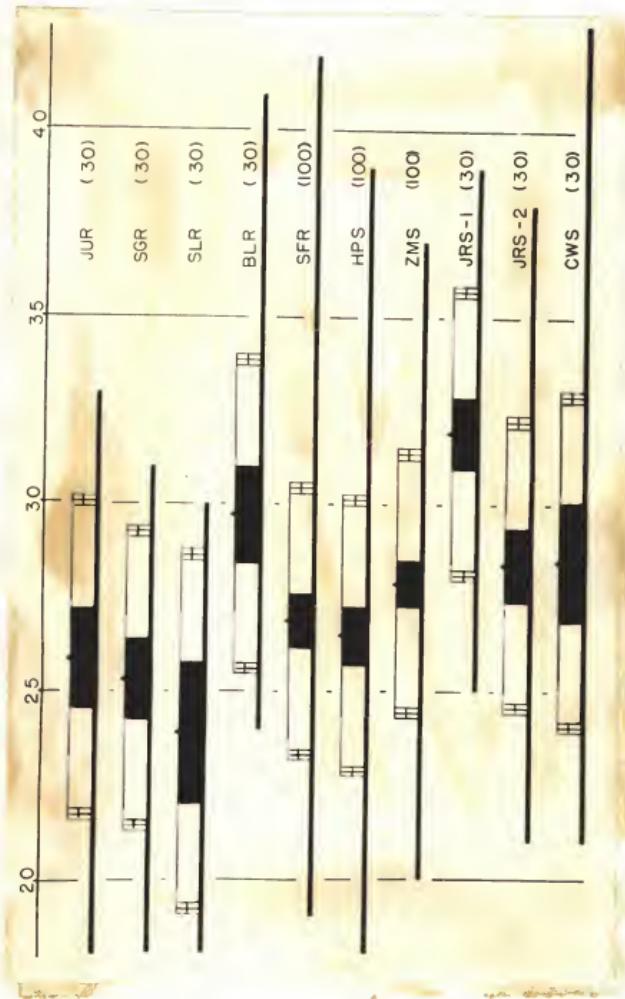


Figure 2. Population-range diagram for angle of gape. The representation is the same as that of Pimentel (1958); i.e., the heavy line indicates the range of variation; the blackened part of bar is the mean (indicated by the small triangle) $\pm s\bar{x}$; the horizontal line inside the open bar on the end represents the 95% confidence interval of the population variance; and distance from the vertical line to the blackened bar represents the pooled sample standard deviation.

MULTIVARIATE ANALYSIS

Mahalanobis' Generalized Distance Function

The multivariate technique chosen for this study is Mahalanobis' Generalized Distance Function (Rao, 1952). This method was chosen over Fisher's Discriminant Function, Hotelling's T^2 , and the Factor Analysis of psychometry largely because of the ease with which one can arrive at an intuitive interpretation of the Distance Function.

The Distance Function has been developed almost exclusively by members of the Indian Statistical Institute of Calcutta. It has been used for distinguishing between neurotic groups (Rao and Slater, 1949) and between phases of locusts (Nair, 1951; Blacklith, 1957), as well as in several anthropological investigations (Mahalanobis, Majumdar, and Rao, 1949; Majumdar and Rao, 1958). Bailey (1956) examined the now classical Partula variation (Crampton, 1916, 1932) by applying this statistic. Although he questioned the effectiveness of genetic drift as a cause, Bailey confirmed Crampton's conclusions as to the distinctness of local populations of this snail.

For exploratory investigations, the computation of the Generalized Distance (D^2) has many advantages over other multivariate techniques. The very fact that Mahalanobis chose the term "distance" for the statistic should make it attractive to students of speciation and racial-

tion. Some kind of distance is certainly involved in the formation of species and races, although this distance, regardless of how it is measured, is of course merely a measurement of the probability of gene flow. That organisms do not recognize the geographic or other distances that men measure is emphasized by Dobzhansky's (1951) definition of the terms "allopatric" and "sympatric."

When a biologist performs a t test, he is in a sense determining the distance between estimates of the means of two populations of single variables; he has converted differences in morphology to a linear scale with standard units, and he asks the question, "Are these estimates of the average values for this character far enough apart on this scale to reject the hypothesis that they estimate the same population mean?" Hubbs' and Hubbs' (1953) and Pimentel's (1958) graphic representations of variability help one to visualize a statistical distance between populations with respect to a single character. With two uncorrelated characters, and with enough different shading techniques, this device could be extended by superimposing the two figures at right angles to each other. Discordant variation would, of course, render the resulting figure less attractive if not meaningless. Further extension to three characters, while difficult to represent visually, is theoretically possible by adding a third dimension.

The Generalized Distance Function also assists in the visualization of differences between populations up to three dimensions. Beyond that, however, graphic representation becomes impossible and the computations become laborious even on an electric desk calculator. However, with the rapidly increasing availability of digital computers (such as the IBM 650), the complexity of the computations should constitute no barrier, even to biologists without mathematical inclination.

Geometric Derivation of D²

In multivariate analysis, geometric interpretation (even for more than three dimensions) is frequently invoked to make clearer the interpretation of results. For example, the Generalized Distance between two populations based on two uncorrelated characters (X_1 and X_2) can be represented by the linear distance between two points plotted on regular Cartesian coordinates. The formula for this

Distance is:

$$D = \sqrt{\frac{(\bar{X}_{11} - \bar{X}_{12})^2}{S_1^2} + \frac{(\bar{X}_{21} - \bar{X}_{22})^2}{S_2^2}}$$

where \bar{X}_{11} and \bar{X}_{12} are the means of X_1 for the first and second populations, respectively; \bar{X}_{21} and \bar{X}_{22} the means of X_2 ; and s_1^2 and s_2^2 the pooled estimates of the variances of X_1 and X_2 . Thus, the use of the Generalized Distance requires the assumption that the variance of each variable is the same in all populations. This assumption may

not always be valid, but Rao (1952) has stated that departures from it are not serious if the measures are at all related and if s^2 is estimated from a large number of measurements. For operations with this statistic, extraction of the square root to obtain D is generally delayed until the last operation, D^2 being used directly. This practice will be followed in the present analysis.

If this X_1 and X_2 are not uncorrelated, Cartesian coordinates are not valid for representation of the distance. In this situation, coordinates are constructed which are not perpendicular but which are inclined to each other at an angle whose cosine is r , the coefficient of correlation between the variables. In the special case of uncorrelated variables, $r = 0$, and the angle whose cosine is zero is 90° , so that the axes are perpendicular. The formula given above must be modified to include any lack of orthogonality due to correlated variables. In algebraic notation:

$$D^2 = \frac{(\bar{X}_{11} - \bar{X}_{12})^2}{s_{12}^2} + \frac{(\bar{X}_{11} - \bar{X}_{12})(\bar{X}_{21} - \bar{X}_{22})}{s_{12}} + \frac{(\bar{X}_{21} - \bar{X}_{22})^2}{s_{22}^2}$$

with the symbols as above except that s_{12} is the covariance of variables X_1 and X_2 .

It is evident that such formulae will become increasingly unwieldy as the number of variables increases, so it is much more convenient to use summation rather than algebraic notation. Letting

$d_1 = (\bar{X}_{11} - \bar{X}_{12})$, $d_2 = (\bar{X}_{21} - \bar{X}_{22})$ --i.e., $d_1 = (\bar{X}_{11} - \bar{X}_{12})$ and s_{ij} be the covariance of the i^{th} and j^{th} variables (or the variance of the i^{th} variable when $i = j$):

$$\underline{D}^2 = \sum (s_{ij})^{-1} d_i d_j.$$

The superscript " -1 " indicates the inverse of the variance-covariance matrix. Both Rao (1952) and Royce (1957) present methods for simplifying the computations of \underline{D}^2 .

Thus, we can see that by scaling Cartesian coordinates in standard deviation units and applying the Pythagorean theorem, the distance between population means of the two independent variables can be estimated; and that this concept can be extended to correlated variables and to three or even more dimensions. Although the computations would be involved and laborious, the use of digital computers and the ease of interpretation make Mahalanobis' Generalized Distance Function an extremely useful statistic for simultaneous analysis of several taxonomic characters. It is also interesting to note at this point that, contrary to Mayr, Linsley, and Usinger's (1953) assertion that the coefficient of difference had not been extended to multiple characters, D^2 is one of several statistically valid measures of differences in multiple characters which antedate their D.D. by about a quarter of a century.

Computation of D^2

Since there is not yet available a program for computing D^2 on the IBM 650, this computation was performed using parts of several other programs. In order to compute also the single factor statistics here reported, the sum of squares and products matrix was determined separately for each sample. Then, because the variances were not the same in all samples and because Rao (1952) suggested that this would not invalidate the technique if the variances and covariances were based on a large number of specimens, the sums of squares and products for all of the 635 specimens were pooled for an estimate of the variance-covariance matrix based on 620 degrees of freedom. This matrix was then inverted; the differences between the means of the 12 measurements for all 45 pairs of samples in the Main Series and Special Series were determined, and the vector-matrix-vector multiplication was performed to arrive at a value of D^2 for each pair of samples. Table 4 presents these values of D^2 . The variance-covariance matrix and its inverse are presented in the Appendix in Tables 10 and 11, respectively.

Interpretation of D^2

If the two subspecies described by Hubbs and Crowe (1956) are valid, and if this divergence is reflected in the morphometric

characters included in this analysis, the populations from which these samples were taken would form two distinct clusters. While the criteria for clusters are not rigidly fixed, "any two groups belonging to the same cluster should, at least on the average, show a smaller D^2 than those belonging to different clusters" (Rao, 1952).

TABLE 4
VALUES OF D^2 , CORRECTED FOR BIAS

	ZMS	SFR -1	HPS	BLR	SGR	SLR	JUR	CWS	JRS -1
SFR-1	6.1								
HPS	2.8	2.0							
BLR	5.5	3.7	3.0						
SGR	5.5	2.4	2.6	33.6					
SLR	33.4	6.4	4.5	37.4	3.4				
JUR	7.0	2.2	5.8	162.9	4.1	7.0			
CWS	4.0	3.8	1.3	157.7	5.0	8.3	9.1		
JRS-1	7.5	7.5	7.2	95.9	20.3	47.5	16.8	16.8	
JRS-2	8.8	9.3	6.9	54.4	33.6	6.4	398.2	392.3	7.4

The first question in the present case is, therefore, whether two clusters can be found, one consisting of the running-water samples and the other comprising the still-water samples. When such clusters are formed, the intracluster average D^2 is 26.31 for the

running water and 45.17 for the standing water. The average intercluster D^2 is 36.92; thus, this arrangement does not fulfill the criterion suggested above.

Since the habitats of the subspecies in the original description are not delimited as above, but rather as "springs and spring-fed creeks" contrasted with "sinks and caves" (Hubbs and Crowe, 1956), it might be necessary to leave SFR-1 and HPS out of consideration and cluster the other samples according to these habitats. With this arrangement, the spring intracluster average D^2 is 41.40; that of the sink hole habitat is 72.27; and the average D^2 between the two habitats is 54.82. This arrangement likewise fails to satisfy the criterion for clustering.

Since neither of these a priori cluster formations will satisfy the criterion, an alternative procedure is to search for clusters in the manner suggested by K. D. Tocher to Rao (1952). Beginning with two samples with a low D^2 , other samples are added to the cluster until such an addition causes a large increase either in the average D^2 or in the average increase in D^2 . The simultaneous development of two clusters sometimes allows a better decision regarding the retention or discarding of a sample in a cluster.

Following this procedure, and beginning the two clusters with ZMS and SFR-1 of the Special Series, the following clusters are

discernible:

A₁, including samples SFR-1, JUR, SGR, and SLR;

A₂, including samples ZMS and CWS;

A₃, including samples JRS-1 and JRS-2; and

A₄, including only sample BLR.

Sample HPS is not included in any of these clusters, because the value of D^2 between this and each of the other samples is so low that HPS can be added to any of the clusters at any time without violating the stipulation for cluster formation. The average intercluster and intracluster values for D^2 are presented in Table 5.

TABLE 5

INTRACLUSTER AND INTERCLUSTER VALUES FOR D^2 IN THE CLUSTERING CONFIGURATION "A"

	A ₁	A ₂	A ₃	A ₄	HPS
A ₁	4.25	9.78	64.38	59.4	4.05
A ₂	9.78	4.0	132.82	81.6	4.70
A ₃	64.38	132.82	7.4	75.15	7.17
A ₄	59.4	81.6	75.15	----	3.0
HPS	4.05	4.70	7.17	3.0	---

Cluster A₁ includes samples SFR-1, JUR, SGR, and SLR

Cluster A₂ includes samples ZMS and CWS

Cluster A₃ includes samples JRS-1 and JRS-2

Cluster A₄ includes only sample BLR

Obviously this configuration does satisfy the criterion of lower average values for the intracluster D^2 ; it is hereafter designated as "A" and is illustrated in Figure 3. Clusters A_1 and A_2 seem to be much closer together than the others, and this suggests that at least one of the members of A_2 might belong in A_1 . Addition of CWS to A_1 increases the average intracluster D^2 to 5.17, which is still considerably below the intercluster values of D^2 .

Thus, an alternative configuration may include CWS with A_1 , and since the D^2 between BLR and ZMS is less than that between JRS-1 and JRS-2 and is very little more than the intracluster average D^2 in A_1 , these two samples are also combined to make a cluster. This configuration is designated "B" and is also illustrated in Figure 3. The intracluster and intercluster values for D^2 are presented in Table 6.

While it was difficult, if not impossible, to discern any pattern in the variation of these twelve characters by the methods of single factor analysis, some sort of structure seems to emerge in either one of the cluster configurations presented above. Considering only the criterion of maximum intercluster and minimum intracluster values for D^2 , the arrangement in "B" seems preferable, but configuration "A" also has some points to recommend it. In the first place, with "A" most of the samples from the stream habitat form a

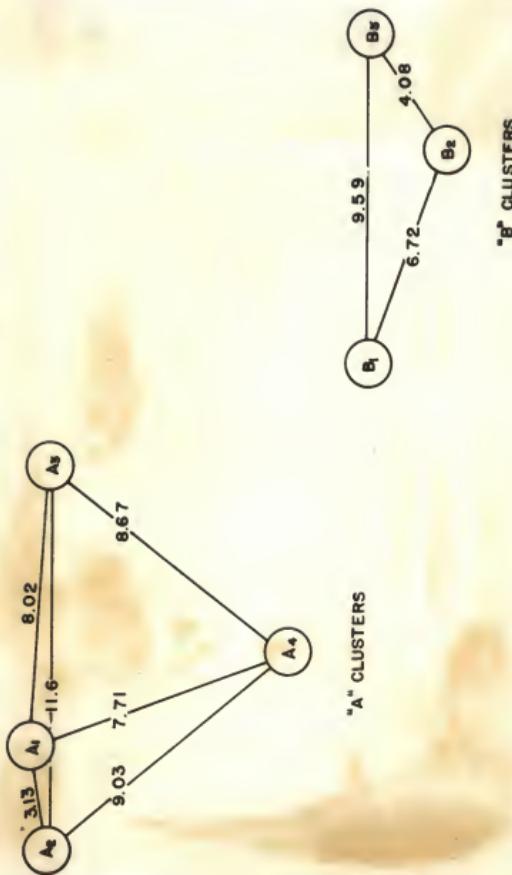


Figure 3. Cluster configurations "A" and "B" with average intercluster values for D .

TABLE 6

INTRACLUSTER AND INTERCLUSTER VALUES FOR D^2 IN THE CLUSTERING CONFIGURATION "B"

	B ₁	B ₂	B ₃	HPS
B ₁	5.17	45.13	91.98	3.26
B ₂	45.13	5.5	16.66	2.85
B ₃	91.98	16.66	7.4	7.05
HPS	3.26	2.85	7.05	----

Cluster B₁ includes samples SFR-1, JUR, SGR, SLR,
and CWS
Cluster B₂ includes samples ZMS and BLR
Cluster B₃ includes samples JRS-1 and JRS-2

distinct, separate cluster. The failure of BLR to be included in this cluster can be explained because this sample was taken in the back-water around the boil of the spring during a period of very high water. Therefore, it may have sampled a population which is less similar to the spring form, either because of actual relationship to the sink hole populations or because of a sorting out in this area of less positively rheotropic individuals.

Also, in "A" it appears that there may be two major axes in the distribution of the clusters: one placed horizontally in the figure and the other running almost vertically near A₁ and A₂. Under a hypothesis of such axes, the horizontal one might be interpreted as

representing variation related to the sink hole habitat and the vertical one as related to the stream habitat. This would thus lead to the conclusion that the range of variability of the sink hole form is so broad that it extends to both sides of the stream-form cluster. This could also account for the small Distances which allow inclusion of CWS in Cluster A₁ and for the fact that HPS can be added to any of the clusters without increasing the intracluster Distance excessively.

Although the configuration in "B" meets the criterion for cluster formation, it is possible that this figure is too greatly oversimplified. This arrangement does emphasize the possibility, mentioned above, that BLR is a sample more representative of the sink hole form than of the spring form. This hypothesis is further supported by the fact that Cluster B₂ is closer to B₃ than it is to B₁.

There is actually little or no contradiction in the two figures. If "A" were to be rotated until BLR were projected onto A₂, a figure very much like "B" would be obtained. The configuration in "A" is to be preferred because this rotation would in effect bring about a greater reduction in the information conveyed by the figure. The entire process of cluster formation that produced "A" resulted in a loss of information, but these rotations were necessary to obtain a structure that could be interpreted and graphically represented.

DISCUSSION AND CONCLUSIONS

In recent years the nonrandomness of samples has come gradually to be seen as a major source of error in taxonomic work. Royce (1957) commented on the difficulty of insuring randomness in samples of fish, and Pimentel (1958) urged systematists to seek complete randomization of collections. A statistical test in which departures from randomness are not serious is much to be desired because most taxonomic work and studies of variation are forced to rely to a large extent on museum collections. Pimentel suggested the application of a test for homogeneity of variances, this to be followed by transformations of the variables until homogeneity was obtained. An alternative procedure is the use of Mahalanobis' D^2 , which can be corrected for bias due to different sample numbers and which, if the estimates of the variances are based on a large number of degrees of freedom, is not greatly affected by departures from homogeneity of variances.

The material on which this study is based is subject to all the departures from randomness to be found in museum collections. Random samples from such widely divergent habitats as caves, sink holes, and springs would be practically impossible to obtain, since the same collecting techniques are not equally effective in these situa-

tions. For example, on a day when fifteen minutes' trapping effort yielded more than 30 fish in Jerome Sink, an hour's trapping in Blue Springs yielded none, even though great numbers of fish could be observed swimming about the trap and bait.

Thus, even though the results must be interpreted with caution because of the heterogeneity of variance, the analysis of variance technique failed to reveal any characters for which variation could be related to habitat distribution. Furthermore, analysis of a series of samples from one area of the Santa Fe River showed that most of these variables are subject to significant sampling variation.

However, when the data were processed through a multivariate technique with D^2 , a structure emerged which is generally in accord with the results hypothesized above from the distribution of Hybopsis. There are, of course, some rather large departures from the predicted results, but these can be explained on the basis of deviations either in the population structure or in the sampling procedure. Some of these deviations might also have affected the single factor analysis, but in the absence of a basic pattern, the deviations could not be identified.

The reduction, by clustering, of the twelve-dimensional hyper-space in which the samples were distributed by the Generalized

Distance analysis results in a distribution of the clusters along two major axes (Figure 3). The horizontal axis in this figure can be taken as representing the range of variation of the sink hole form of Hybopsis harperi. Hypothetically, a single isolate from a sink hole population would produce a subpopulation from which a sample could be found somewhere along this axis. A second, less well-defined axis can be distinguished, and this one may represent the dimension through which the stream form varies. On this axis, BLR, a sample from a part of a spring with much reduced current, is found near one extreme. The other stream samples form a distinct cluster close to the intersection of this axis and the sink hole axis.

To continue this geometric interpretation: we can visualize a change in morphometric characters, regardless of cause, to be manifested as movement of a population through a hyperspace which is represented here by only two or three dimensions. The characteristics of the stream habitat produce movement, however slight, mainly in the direction of the A₁-A₄ axis, and the characteristics of the sink hole habitat produce changes along the A₂-A₃ axis. Since sink hole populations are almost certainly derived ultimately from stream populations, the sink hole samples may be considered to have moved away from the cluster of stream populations along the axis of the sink hole habitat.

The extreme position of BLR on the stream habitat axis has already been mentioned and a possible explanation suggested. There are several other points which require further explanation: the location of CWS which would allow its inclusion in the stream sample cluster (A_1) and the very large distance between this sample and JRS-2, while its Distance from JRS-1 is much less; the small Distances between HPS and all the other samples; and the relatively large Distance between JRS-1 and JRS-2.

From 1955 until 1959, the general water table was low, and in Cow Sink only about fifteen square feet of water surface was exposed. During this time the population of Hybopsis was very small. In fact, on several occasions one could see only two or three fish in the sink hole. Throughout this time, Cow Sink was used as a source of irrigation water during the tobacco growing season. In 1959 the water level rose considerably, and the Hybopsis population of Cow Sink increased greatly. Two factors in this situation might explain the small distance between CWS and the cluster of stream samples. First the environment is, with respect to the fundamental factor of water flow, comparable to a small intermittent spring with a flow of about 10,000 gallons per hour (the output during the irrigation season). Thus, some environmentally produced phenotypic effect might hold CWS nearer the intersection of the two main axes. On the other

hand, the population which is sampled by CWS might be derived either from the very reduced population that survived the period of low water or from a colonization by new migrants into the sink hole. In either case, the genetic composition of the newly increased population would be similar to that of the stream population. Obviously, new migrants would be expected to bring a random sample of the genes of the population from which they came. And if, as Lerner (1954) has suggested, "natural selection will tend to favor organisms clustering around mean values for all characters," the survivors of the period of reduced population would be expected to have a genetic constitution like that of the average population, which, according to Figure 3, is the stream form. In any case, even though CWS is included in the cluster with ZMS, it is much closer to the stream-form cluster than is ZMS.

The peculiar relationship of the Distances between JRS-1, JRS-2, and CWS is also related to the physical changes described above, as is the relatively large Distance between JRS-1 and JRS-2. Beginning around 1956, Jerome Sink also was used for irrigation. Therefore, the sample taken in 1959 (JRS-1) was from a population subject to the same conditions as those described above for CWS, except that at no time was the population of Jerome Sink so drastically reduced. This population may have thus moved along

the sink hole axis back toward the stream-form cluster after irrigation began. By the same token, irrigation between the times when JRS-2 and JRS-1 were collected would have reduced the effectiveness of isolation and have allowed the Distance between the population samples by CWS and that sampled by JRS-1 to decrease.

The situation of HPS is more difficult to explain. The greatest Distance between HPS and any other sample is that between HPS and JRS-1 ($D^2 = 7.2$); this value is less than that between JRS-1 and JRS-2. The population sampled by HPS definitely represents a new isolate because the pond was formed by scooping out a mud hole in the bottom of the lime rock mine. Either by influx of many migrants, or by the very rapid reproduction of a few, the population increased to an estimated several thousand in a period of two years. Again it is possible to invoke the idea of natural selection favoring the average to account for the small distance between HPS and the stream-form cluster, but neither this mechanism nor the hypothesis of large scale immigration will explain the small distance between HPS and the sink hole samples. In fact, the only hypothesis that seems to explain the location of HPS is that the movement through the hyperspace was not limited to either of the two major axes herein described. Nor was it limited to movement along both these axes. The movement of this population was apparently in some direction

such that projection to these axes places HPS near each of the clusters. This condition could be realized if the aspects of this habitat that are similar to the stream habitat (e.g., direct sunlight, high basic productivity) produce movement that can be referred to that axis, and if those aspects similar to the sink hole habitat (e.g., no current, no submerged or emergent tracheophytes) produce movement that can be referred to that axis.

Thus, the use of Mahalanobis' D^2 and the clustering technique allow formation of an array of samples from the various habitats that is meaningful because it raises fundamental questions concerning the genetic and ecological relationships among the populations. Moreover, this interpretation of D^2 provides a model for the morphometric divergence of subpopulations which has previously been interpreted as subspeciation.

A final conclusion to be drawn from this analysis is that the structure of the clusters of samples does not warrant recognition of two subspecies in any conventional taxonomic sense. The populations inhabiting springs and spring-fed streams are similar enough to each other for consideration as a single, discrete taxonomic entity. On the other hand, the divergence of the sink hole populations is not consistently in one direction, as one would expect

it to be if there were two and only two subspecies of Hybopsis harperi. The alternatives that present themselves are: (1) to name and describe several subspecies from sink holes, producing even more discontinuities and interdigitations in maps of the ranges of the subspecies, or (2) simply to describe H. harperi as a variable species, in which local isolates diverge from the typical stream form in at least two directions. I prefer the latter course of action.

SUMMARY

Single factor analysis for five meristic characters of 450 fish in eight samples of Hybopsis harperi reveal little variation in these characters in the various subpopulations. The character most strongly supporting the validity of the subspecies now recognized taxonomically is the number of scales above the lateral line.

Similar analysis of eight linear measurements and four ratios for those 450 fish and for 90 more in three other samples failed to show any pattern of variation that could be related to the habitat. Variation in the angle of gape was also studied; and although there seemed to be a consistent difference between samples from the two different kinds of habitat, the results of graphic analysis of this character and the questionable precision of the technique of measurement preclude its use as a diagnostic subspecific character in H. harperi.

To investigate sampling variation, the measurements, ratios, and angle were analyzed in a series of samples from the same location. This analysis indicated that, with the possible exception of the angle, any one of these characters is likely to show: (1) differences in samples collected by different methods; and (2) differences due to unidentified causes.

After it was demonstrated that single factor analysis of the

material available for this study did not reveal any characters with concordant variation, the measurements and ratios were subjected to multivariate analysis with Mahalanobis' D^2 . Through use of this technique variational structure does emerge, and the estimated Distances between the various populations permit the following conclusions concerning the relationship between the populations:

1. In the twelve-dimensional hyperspace formed by this analysis, there is a major axis, along which the sink hole populations vary from each other.
2. Intersecting this axis there may be another which represents the range of variation of the spring and stream populations.
3. At or near the point of intersection of these axes, the stream populations form a distinct cluster.
4. A model for morphometric change is suggested, in which change is interpreted as movement along the axis corresponding to the habitat of an isolated population.
5. Departures from the above pattern can be explained on the basis of:
 - (a) changes in the physical environment,
 - (b) rapid increase in population numbers which allows survival of extreme variants, or
 - (c) selection for the average during periods of lowered

water level and reduced population (genetic homeostasis).

6. The failure of the sink hole populations to form one distinct cluster and the fact that movement along the "sink hole axis" may be in either direction seem to be grounds for concluding that the taxonomic separation of Hybopsis harperi into two, and only two, subspecies does not satisfactorily reflect the natural situation.

LITERATURE CITED

- Bailey, D. W. 1956. Re-examination of the diversity in Partula taeniata. *Evolution*, 10:360-366.
- Blackith, R. E. 1957. Polymorphism in some Australian locusts and grasshoppers. *Biometrics*, 13:183-196.
- Borgmeier, T. 1957. Basic questions of systematics. *Syst. Zool.*, 6:53-69.
- Brown, W. L., Jr., and Wilson, E. O. 1954. The case against the trinomen. *Syst. Zool.*, 3:174-175.
- Cochran, W. G. 1947. Some consequences when the assumptions for the analysis of variance are not satisfied. *Biometrics*, 3:22-38.
- Cochran, W. G., and Cox, G. M. 1957. Experimental designs (Second edition). John Wiley and Sons, Inc., New York, 611pp.
- Crampton, H. E. 1916. Studies on the variation, distribution, and evolution of the genus Partula. The species inhabiting Tahiti. Carnegie Inst. Washington Publ. 228:1-311.
- _____. 1932. Studies on the variation, distribution, and evolution of the genus Partula. The species inhabiting Moorea. Carnegie Inst. Washington Publ. 410:1-335.
- Dobzhansky, T. 1951. Genetics and the origin of species (Third edition). Columbia University Press, New York, 364 pp.

- Edwards, J. G. 1956. Clarification of certain aspects of infra-specific systematics. *Syst. Zool.*, 5:92-94.
- Gosline, W. A. 1954. Further thoughts on subspecies and trinomials. *Syst. Zool.*, 3:92-94.
- Hellier, T. R. 1957. The fishes of the Santa Fe River System. M. S. Thesis, Library, University of Florida.
- Hubbell, T. H. 1954. The naming of geographically variant populations or what is all the shooting about? *Syst. Zool.*, 3:113-121.
- Hubbs, C. L., and Crowe, W. R. 1956. Preliminary analysis of the American cyprinid fishes, seven new, referred to the genus Hybopsis, subgenus Erimystax. *Occ. Papers Mus. Zoo., Univ. Mich.*, 578:1-8.
- Hubbs, C. L., and Hubbs, C. 1953. An improved graphical analysis and comparison of series of samples. *Syst. Zool.*, 2:49-56.
- Hubbs, C. L., and Lagler, K. F. 1941. Guide to the fishes of the Great Lakes and tributary waters. *Bull. Cranbrook Inst. Sci.*, 18:1-100.
- Lerner, I. M. 1954. Genetic Homeostasis. John Wiley and Sons Inc., New York, 134 pp.
- Li, J. C. R. 1957. Introduction to statistical inference. Edwards Bros., Ann Arbor, 553 pp.

- Mahalanobis, P. C., Majumdar, D. N., and Rao, C. R. 1949. Anthropometric survey of the United Provinces, 1941: A statistical study. *Sankhya*, 9:20-266.
- Majumdar, D. N., and Rao, C. R. 1958. Bengal anthropometric survey, 1945: A statistical study. *Sankhya*, 19:201-408.
- Marr, J. C. 1957. The problem of defining and recognizing subpopulations of fishes. U. S. Fish and Wildlife Service Special Sci. Report, 208:1-6.
- Marshall, N. 1941. A study of the life history and ecology of Notropis chalybaeus (Cope) supplemented with data on other cyprinids in Florida. Ph. D. Thesis, Library, University of Florida.
- _____. 1947. The spring run and cave habitats of Erimystax harperi (Fowler). *Ecology*, 28:68-75.
- Mayr, E., Linsley, E. G., and Usinger, R. L. 1953. Methods and principles of systematic zoology. McGraw-Hill, New York, 328 pp.
- Nair, K. R. 1951. A biometric study of the desert locust. Proc. Indian Statistical Conference, 1951, 349-358.
- Olson, E. C., and Miller, R. L. 1958. Morphological Integration. University of Chicago Press, Chicago, 317 pp.
- Ostle, B. 1954. Statistics in Research. Iowa State College Press, Ames, Iowa, 487 pp.
- Pimentel, R. A. 1958. Taxonomic methods, their bearing on speciation. *Syst. Zool.*, 7:139-156.

- Rao, C. R. 1952. Advanced statistical methods in biometric research. John Wiley and Sons, New York, 390 pp.
- Rao, C. R., and Slater, P. L. 1949. Multivariate analysis applied to differences between neurotic groups. British Jour. Psychol., Statistics Sect., 2:17-32.
- Royce, W. F. 1957. Statistical comparison of morphological data. U. S. Fish and Wildlife Service Special Scientific Report, 208:7-28.
- Snedecor, G. W. 1956. Statistical methods applied to experiments in agriculture and biology. Iowa State College Press, Ames, Iowa, 534 pp.
- Wilson, E. O., and Brown, W. L., Jr. 1953. The subspecies concept and its application. Syst. Zool., 2:97-111.
- Wright, S. 1931. Evolution in Mendelian populations. Genetics, 16:97-159.

APPENDIX

This appendix contains the complete statistical analyses referred to in the text. In most instances, the entries in the tables contain fewer digits than were actually used in the computations. For calculating means, sums of squares and products, and standard deviations, ten digits were used. For Table 7, the means and standard deviations of the measurements were converted to millimeters.

The matrix operations involved in computing D^2 were performed in floating point arithmetic; therefore, the numbers actually used contained eight digits in scientific notation. For example:

$$31,768.453 = 3.1768453 \times 10^4.$$

TABLE 7

MEANS AND STANDARD DEVIATIONS
FOR ALL CHARACTERS IN ALL SAMPLES

	AG	L	DB	HL	SL	E	LD
ZMS	27.8 (4.00)	31.6 (6.45)	6.9 (1.52)	7.9 (1.43)	2.1 (0.50)	2.3 (0.31)	16.2 (3.33)
HPS	26.5 (3.74)	31.9 (4.19)	8.1 (1.02)	8.0 (1.03)	2.2 (0.34)	2.2 (0.23)	16.1 (2.26)
SFR-1	27.9 (3.46)	37.6 (4.82)	8.0 (1.02)	8.6 (1.01)	2.4 (0.34)	2.7 (0.24)	19.2 (2.53)
SFR-2	26.6 (3.25)	37.5 (6.20)	8.8 (1.33)	8.2 (1.34)	2.2 (0.41)	2.5 (0.30)	18.7 (3.32)
SFR-3	29.7 (4.00)	33.6 (3.33)	8.0 (0.97)	7.1 (0.63)	1.9 (0.25)	2.2 (0.16)	17.0 (1.68)
SFR-4	28.9 (3.25)	32.9 (2.59)	7.3 (0.86)	7.0 (0.63)	1.8 (0.25)	2.2 (0.15)	16.0 (1.41)
SFR-5	28.1 (2.69)	31.3 (3.19)	6.6 (0.72)	7.0 (0.79)	1.8 (0.23)	2.3 (0.22)	15.3 (1.65)
SFR-6	27.5 (2.36)	38.1 (3.09)	8.6 (0.91)	9.6 (0.76)	2.2 (0.16)	2.4 (0.27)	18.7 (1.86)
JRS-1	31.8 (5.47)	34.6 (7.44)	8.5 (1.89)	7.7 (1.53)	1.9 (0.48)	2.3 (0.34)	18.5 (3.80)
JRS-2	28.4 (3.46)	23.6 (3.42)	5.3 (0.79)	5.5 (0.75)	1.3 (0.25)	1.7 (0.16)	12.5 (1.62)
CWS	28.5 (5.57)	36.3 (7.79)	8.0 (2.31)	8.7 (1.78)	2.3 (0.58)	2.5 (0.31)	18.8 (4.46)
BLR	29.7 (2.83)	28.4 (4.86)	6.2 (1.02)	6.3 (0.98)	1.6 (0.33)	2.1 (0.24)	14.6 (2.51)
SGR	25.3 (3.00)	32.5 (4.94)	7.0 (1.17)	7.4 (0.95)	1.8 (0.27)	2.3 (0.19)	15.9 (2.46)
SLR	23.9 (3.00)	24.4 (6.02)	5.4 (1.14)	5.7 (0.35)	1.4 (0.35)	1.9 (0.27)	12.4 (2.88)
JUR	25.9 (3.60)	36.8 (5.74)	6.9 (1.17)	8.2 (1.17)	2.3 (0.41)	2.6 (0.29)	18.5 (2.84)

TABLE 7--Continued

	PD	HW	HL/SL	HL/E	HL/HW	PD/DB
ZMS	2.5 (0.49)	4.2 (0.75)	3.85 (0.41)	3.35 (0.28)	1.91 (0.12)	0.37 (0.04)
HPS	2.7 (0.34)	4.4 (0.57)	3.73 (0.25)	3.54 (0.19)	1.81 (0.12)	0.33 (0.03)
SFR-1	2.9 (0.40)	4.6 (0.60)	3.66 (0.20)	3.21 (0.21)	1.98 (0.08)	0.36 (0.03)
SFR-2	2.9 (0.47)	4.4 (0.71)	3.71 (0.27)	3.24 (0.25)	1.85 (0.06)	0.33 (0.02)
SFR-3	2.8 (0.29)	4.0 (0.43)	3.86 (0.30)	3.16 (0.15)	1.80 (0.08)	0.36 (0.02)
SFR-4	2.7 (0.24)	3.8 (0.34)	3.95 (0.23)	3.18 (0.20)	1.85 (0.07)	0.36 (0.02)
SFR-5	2.6 (0.34)	3.7 (0.40)	3.93 (0.19)	3.10 (0.16)	1.90 (0.06)	0.39 (0.03)
SFR-6	3.1 (0.27)	4.5 (0.35)	3.97 (0.25)	3.55 (0.25)	1.90 (0.08)	0.37 (0.03)
JRS-1	2.6 (0.53)	4.3 (0.96)	4.06 (0.39)	3.36 (0.33)	1.80 (0.14)	0.32 (0.03)
JRS-2	1.8 (0.29)	3.0 (0.41)	4.15 (0.34)	3.14 (0.24)	1.83 (0.08)	0.34 (0.02)
CWS	2.8 (0.62)	4.6 (1.07)	3.80 (0.38)	3.49 (0.39)	1.82 (0.13)	0.36 (0.04)
BLR	2.2 (0.36)	3.5 (0.58)	4.01 (0.36)	3.07 (0.17)	1.83 (0.07)	0.36 (0.02)
SGR	2.5 (0.35)	3.9 (0.56)	4.05 (0.19)	3.18 (0.20)	1.91 (0.09)	0.36 (0.02)
SLR	2.0 (0.46)	3.0 (0.61)	4.24 (0.30)	3.06 (0.20)	1.95 (0.07)	0.36 (0.02)
JUR	2.9 (0.41)	4.3 (0.57)	3.59 (0.25)	3.18 (0.25)	1.91 (0.12)	0.34 (0.02)

TABLE 7--Continued

	D	A	LL	SA	SB
ZMS	8.00 (0.14)	7.88 (0.38)	35.8 (1.48)	5.21 (0.43)	2.99 (0.27)
HPS	7.99 (0.22)	8.07 (0.38)	36.7 (1.01)	5.41 (0.49)	3.11 (0.32)
SFR-1	8.02 (0.14)	8.03 (0.17)	36.7 (0.99)	5.92 (0.27)	3.04 (0.40)
JRS-1	7.93 (0.25)	8.10 (0.30)	36.5 (1.31)	5.77 (0.43)	3.17 (0.38)
JRS-2	8.00 (0.00)	8.07 (0.25)	35.1 (1.09)	5.63 (0.49)	3.00 (0.37)
CWS	8.07 (0.36)	8.07 (0.25)	36.7 (0.90)	5.70 (0.54)	3.07 (0.45)
BLR	8.00 (0.00)	7.97 (0.18)	36.1 (1.07)	6.00 (0.26)	3.13 (0.35)
JUR	8.00 (0.00)	8.03 (0.32)	36.5 (1.14)	5.93 (0.25)	3.07 (0.36)

TABLE 8
VALUES OF CHI SQUARE FOR BARTLETT'S TEST

Character	Special Series	Santa Fe Series	Main Series
D	30.27***	--	--
A	67.10***	--	10.28*
LL	21.91***	--	3.69
SA	34.03***	--	28.56***
SB	17.04***	--	2.48
AG	2.08	13.71**	5.86
L	19.66***	37.03***	32.24***
DB	21.93***	23.52***	58.14***
HL	15.71***	30.23***	35.94***
SL	20.78***	20.80***	38.27***
E	11.88**	30.31***	31.00***
LD	16.19***	35.78***	41.63***
PD	12.66**	23.66***	164.28***
HW	8.57*	30.62***	46.23***
HL/SL	54.32***	11.64*	26.03***
HL/E	16.23***	149.63***	37.02***
HL/HW	16.79***	11.18*	37.47***
PD/DB	5.53	21.49***	26.07***

* P ≤ 0.05

** P ≤ 0.01

*** P ≤ 0.001

TABLE 9
COEFFICIENTS AND DIVISORS FOR COMPUTING
SUMS OF SQUARES FOR ORTHOGONAL COMPARISONS

Samples	Coefficients											
	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈	C ₉	C ₁₀ ^b	C ₁₁	C ₁₂
ZMS	-1	-1										
HPS	-1	+1										
SFR-1	+2											
SFR-2			+10									
SFR-3			+12									
SFR-4			-15	+1								
SFR-5			-15	-1								
SFR-6			+10			-1						
JRS-1					+4		-1		+1			
JRS-2					+4		-1		-1			
CWS					+4		+1					
BLR					-3				+3			
SGR					-3				-1	-1	+1	
SLR					-3				-1	-1	-1	
JUR					-3				-1	+2		
Divisors	600	200	23100	60	60	-	2520	180	60	360	180	60

a/ The sum of squares for C₆ was obtained by subtraction (Rule 4, Cochran and Cox, 1957).

b/ For the meristic characters, +1 (BLR) and -1 (JUR) were used with the divisor = 60.

TABLE 10

VARIANCE-COVARIANCE MATRIX FOR MEASUREMENTS AND RATIOS

	L	DB	HL	SL	E	LD	PD	HW	HL/SL	HL/E	HL/HW	PD/DB
L	411.4											
DB	88.6	24.5										
HL	752.6	169.3	1628.8									
SL	240.0	54.7	514.0	181.7								
E	173.6	36.2	348.0	109.3	91.5							
LD	203.7	45.3	382.4	122.2	87.4	108.7						
PD	277.9	64.0	524.2	166.4	117.1	138.1	209.2					
HW	419.0	98.9	857.6	274.5	184.4	214.4	296.4	500.7				
HL/SL	-331.7	-80.2	-672.9	-337.8	-142.1	-167.0	-321.1	-379.7	1190.6			
HL/E	285.7	79.3	750.4	240.5	63.9	148.1	220.2	390.3	-348.0	826.4		
HL/HW	-24.2	-12.3	22.1	1.7	2.7	-14.2	-24.2	-62.1	30.0	18.7	116.4	
PD/DB	-17.7	-58.8	-64.6	-32.3	7.6	-24.9	21.2	-73.3	116.2	-141.4	57.6	915.7

TABLE 11
INVERSE OF VARIANCE-COVARIANCE MATRIX^{a/}

L	<u>5.95</u>	L	DB	HL	SL	E	LD	PD	HW	<u>HL</u> <u>SL</u>	<u>HL</u> <u>E</u>	<u>HL</u> <u>HW</u>	<u>PD</u> <u>DB</u>
DB	-0.72	485.59											
HL	-0.66	20.08	10.53										
SL	0.56	-17.72	2.96	54.42									
E	-0.11	-6.00	11.29	-14.44	-40.65								
LD	-5.63	-4.47	0.38	-2.52	-0.24	14.48							
PD	-2.83	-144.20	-6.45	8.02	2.54	1.21	49.04						
HW	-0.03	-28.16	-26.69	-24.53	8.76	-0.88	6.58	58.54					
HL/SL	0.12	-2.79	0.64	7.59	-2.30	-0.37	1.23	-3.64	1.17				
HL/E	0.28	-1.50	2.15	-2.70	-8.25	0.11	0.40	1.51	-0.42	-1.42			
HL/HW	-0.09	-13.36	-17.09	-15.81	5.83	0.02	3.74	36.53	-2.38	1.03	24.10		
PD/DB	-0.01	33.94	1.38	-1.48	-0.48	-0.10	-10.42	-1.72	-0.24	-0.08	-0.89	2.51	-2

a/ All values $\times 10^{-2}$

TABLE 12
ANALYSIS OF VARIANCE FOR DORSAL RAYS

Series	Source of Variation	D.F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	0.05	0.025	0.83
	C ₁	1	0.042	0.042	1.38
	C ₂	1	0.005	0.005	0.17
	Within Samples	297	8.95	0.030	
	Total	299	9.00		
Main	Between Samples	4	0.267	0.067	0.17
	C ₇	1	0.2	0.2	0.51
	C ₈	1	0.067	0.067	0.17
	C ₉	1	0.17		
	C ₁₀	1	5.733	0.395	
	Within Samples	145	6.000		
	Total	149			

* P ≤ 0.05

** P ≤ 0.01

*** P ≤ 0.001

TABLE 13
ANALYSIS OF VARIANCE FOR ANAL RAYS

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	2.01	1.00	9.14
	C ₁	1	0.20	0.20	1.87
	C ₂	1	1.81	1.81	16.76***
	Within Samples	297	31.98	0.11	
	Total	299	33.99		
Main	Between Samples	4	0.30	0.08	1.14
	C ₇	1	0.02	0.02	0.30
	C ₈	1	0.01	0.01	0.08
	C ₉	1	0.17	0.17	0.23
	C ₁₀	1	0.07	0.07	0.93
	Within Samples	145	10.37	0.07	
	Total	149	10.67	0.07	

* P ≤ 0.05

** P ≤ 0.01

*** P ≤ 0.001

TABLE 14
ANALYSIS OF VARIANCE FOR SCALES IN LATERAL LINE

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	58.41	29.20	20.96***
	C ₁	1	12.33	12.33	8.85**
	C ₂	1	46.08	46.08	33.08***
	Within Samples	297	413.78	1.39	
	Total	299	472.19		
Main	Between Samples	4	49.37	12.34	9.98***
	C ₇	1	1.28	1.28	1.04
	C ₈	1	15.60	15.60	12.62***
	C ₉	1	30.82	30.82	24.91***
	C ₁₀	1	1.67	1.67	1.35
	Within Samples	145	179.36	1.24	
	Total	149	228.73		

* P ≤ 0.05

** P ≤ 0.01

*** P ≤ 0.001

TABLE 15
ANALYSIS OF VARIANCE FOR SCALES ABOVE LATERAL LINE

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	26.81	13.40	79.40***
	C ₁	1	24.81	24.81	146.94**
	C ₂	1	2.00	2.00	11.85***
	Within Samples	297	50.14	0.17	
	Total	299	76.95		
Main	Between Samples	4	2.89	0.72	4.28*
	C ₇	1	2.56	2.56	15.51***
	C ₈	1	0.00	0.00	0.00
	C ₉	1	0.27	0.27	1.58
	C ₁₀	1	0.06	0.06	0.40
	Within Samples	145	24.50	0.17	
	Total	149	27.39		

* P ≤ 0.05

** P ≤ 0.01

*** P ≤ 0.001

TABLE 16
ANALYSIS OF VARIANCE FOR SCALES BELOW LATERAL LINE

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	0.73	0.36	3.31
	C ₁	1	0.01	0.01	0.06
	C ₂	1	0.72	0.72	6.56*
	Within Samples	297	32.62	0.11	
	Total	299	33.35		
Main	Between Samples	4	0.51	0.12	0.86
	C ₇	1	0.02	0.02	0.12
	C ₈	1	0.01	0.01	0.04
	C ₉	1	0.42	0.42	2.83
	C ₁₀	1	0.06	0.06	0.45
	Within Samples	145	21.36	0.15	
	Total	149	21.87		

* P ≤ 0.05

** P ≤ 0.01

*** P ≤ 0.001

TABLE 17
ANALYSIS OF VARIANCE FOR ANGLE OF GAPE

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	118.00	59.00	4.00*
	C ₁	1	35.00	35.00	2.38
	C ₂	1	83.00	83.00	5.64*
	Within Samples	297	4381.00	14.75	
	Total	299	4499.00		
Santa Fe	Between Samples	4	15795.00	3948.75	3.90**
	C ₃	1	1119.36	1119.36	1.10
	C ₄	1	1083.75	1083.75	1.07
	C ₅	1	1170.42	1170.42	1.16
	C ₆	1	12421.47	12421.47	
	Within Samples	140	141856.00	1013.26	
	Total	144	157657.00		
Main	Between Samples	6	1347.00	224.50	15.84***
	C ₇	1	580.03	580.03	40.93***
	C ₈	1	51.20	51.20	3.61
	C ₉	1	166.67	166.67	11.76***
	C ₁₀	1	487.67	487.67	31.41***
	C ₁₁	1	32.94	32.94	2.32
	C ₁₂	1	28.02	28.02	1.98
	Within Samples	203	2877.00	14.17	
	Total	209	4224.00		

TABLE 18
ANALYSIS OF VARIANCE FOR LENGTH

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	21991.00	10995.50	414.39***
	C ₁	1	21961.50	21961.50	827.66***
	C ₂	1	28.88	28.88	1.09
	Within Samples	297	77807.00	26.53	
	Total	299	100798.00		
Santa Fe	Between Samples	4	10061.00	2515.25	16.42***
	C ₃	1	6142.67	6142.67	40.14***
	C ₄	1	365.07	365.07	2.39
	C ₅	1	46.82	46.82	0.31
	C ₆	1	3506.44	3506.44	22.91***
	Within Samples	140	21284.00	153.03	
	Total	144	31345.00		
Main	Between Samples	6	52416.00	8736.00	25.22***
	C ₇	1	512.10	512.10	11.96***
	C ₈	1	10050.13	10050.13	234.69***
	C ₉	1	17374.02	17374.02	405.72***
	C ₁₀	1	1707.38	1707.38	39.87***
	C ₁₁	1	13347.22	13347.22	311.68***
	C ₁₂	1	9425.07	9425.07	220.09***
	Within Samples	203	70331.00	346.46	
	Total	209	122747.00		

TABLE 19
ANALYSIS OF VARIANCE FOR DEPTH OF BODY

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	915.00	457.50	31.22***
	C ₁	1	189.28	189.28	12.91***
	C ₂	1	725.80	725.80	49.52***
	Within Samples	297	4353.00	14.66	
	Total	299	5268.00		
Santa Fe	Between Samples	4	980.00	245.00	24.60***
	C ₃	1	776.78	776.78	77.99***
	C ₄	1	72.60	72.60	7.29**
	C ₅	1	6.67	6.67	0.67
	C ₆	1	123.95	123.95	12.40***
	Within Samples	140	1395.00	9.96	
	Total	144	2375.00		
Main	Between Samples	6	3286.00	547.67	25.65***
	C ₇	1	100.00	100.00	4.68*
	C ₈	1	235.75	235.75	11.04**
	C ₉	1	1401.67	1401.67	65.66***
	C ₁₀	1	136.90	136.90	6.41*
	C ₁₁	1	1036.80	1036.80	48.56***
	C ₁₂	1	375.00	375.00	17.56***
	Within Samples	203	4334.00	21.35	
	Total	209	7620.00		

TABLE 20
ANALYSIS OF VARIANCE FOR HEAD LENGTH

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	23078.00	11539.00	10.65***
	C ₁	1	22903.08	22903.08	21.14***
	C ₂	1	174.84	174.84	0.16
	Within Samples	297	321697.00	1083.16	
	Total	299	344775.00		
Santa Fe	Between Samples	4	49694.00	12423.50	19.87***
	C ₃	1	23850.00	23850.00	38.15***
	C ₄	1	8.07	8.07	0.01
	C ₅	1	1152.82	1152.82	1.84
	C ₆	1	24683.12	24683.12	39.48***
	Within Samples	140	87519.00	625.14	
	Total	144	137713.00		
Main	Between Samples	6	219728.00	36621.00	29.94***
	C ₇	1	5992.40	5992.40	4,90*
	C ₈	1	73406.80	73406.80	60.02***
	C ₉	1	56242.82	56242.82	45.98***
	C ₁₀	1	10605.88	10605.88	8.67**
	C ₁₁	1	41648.02	41648.02	34.05***
	C ₁₂	1	31832.07	31832.07	26.02***
	Within Samples	203	248294.00	1223.12	
	Total	209	468022.00		

TABLE 21
ANALYSIS OF VARIANCE FOR SNOUT LENGTH

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	3190.00	1595.00	12.65***
	C ₁	1	3064.56	3064.56	24.31***
	C ₂	1	124.82	124.82	0.99
	Within Samples	297	37446.00	126.08	
	Total	299	40636.00		
Santa Fe	Between Samples	4	4218.00	1054.50	15.24***
	C ₃	1	2335.46	2335.46	33.76***
	C ₄	1	0.15	0.15	0.002
	C ₅	1	58.02	58.02	0.84
	C ₆	1	1824.37	1824.37	26.38***
	Within Samples	140	9684.00	69.17	
	Total	144	13902.00		
Main	Between Samples	6	23242.00	3783.67	29.45***
	C ₇	1	325.00	325.00	2.53
	C ₈	1	7973.35	7973.35	62.07***
	C ₉	1	3969.07	3969.07	30.90***
	C ₁₀	1	963.67	963.67	7.50**
	C ₁₁	1	7592.00	7592.00	59.10***
	Within Samples	203	26078.00	128.46	
	Total	209	49320.00		

TABLE 22
ANALYSIS OF VARIANCE FOR WIDTH OF EYE

Series	Source of Variation	D.F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	8059.00	4029.50	73.34***
	C ₁	1	7697.00	7697.00	140.08***
	C ₂	1	361.80	361.80	6.58*
	Within Samples	297	16319.00	54.95	
	Total	299	24378.00		
Santa Fe	Between Samples	4	1632.00	408.00	11.63***
	C ₃	1	694.72	694.72	19.80***
	C ₄	1	24.07	24.07	0.69
	C ₅	1	166.67	166.67	4.75*
	C ₆	1	746.54	746.54	21.28***
	Within Samples	140	4911.00	35.06	
	Total	144	6543.00		
Main	Between Samples	6	13842.00	2307.00	40.25***
	C ₇	1	50.85	50.85	0.89
	C ₈	1	3600.14	3600.14	62.81***
	C ₉	1	3360.02	3360.02	58.62***
	C ₁₀	1	677.88	677.88	11.83***
	C ₁₁	1	3708.27	3708.27	64.69***
	C ₁₂	1	2444.82	2444.82	42.65***
	Within Samples	203	11636.00	57.32	
	Total	209	25478.00		

TABLE 2.3
ANALYSIS OF VARIANCE FOR LENGTH TO DORSAL FIN

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	5875.00	2937.50	40.22***
	C ₁	1	5871.88	5871.88	80.39***
	C ₂	1	2.20	2.20	0.03
	Within Samples	297	21693.00	73.04	
	Total	299	27568.00		
Santa Fe	Between Samples	4	2676.00	669.00	15.06***
	C ₃	1	2031.87	2031.87	45.74***
	C ₄	1	62.01	62.01	1.40
	C ₅	1	0.02	0.02	0.0004
	C ₆	1	582.10	582.10	13.10***
	Within Samples	140	6219.00	44.42	
	Total	144	8895.00		
Main	Between Samples	6	13611.00	2268.50	24.36***
	C ₇	1	751.33	751.33	8.07**
	C ₈	1	2128.67	2128.67	22.84***
	C ₉	1	5208.02	5208.02	55.92***
	C ₁₀	1	217.78	217.78	2.34
	C ₁₁	1	3555.56	3555.56	38.18***
	C ₁₂	1	1749.60	1749.60	18.79***
	Within Samples	203	18905.00	93.13	
	Total	209	32516.00		

TABLE 24

ANALYSIS OF VARIANCE FOR DEPTH OF CAUDAL PEDUNCLE

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	7335.00	3667.50	35.17***
	C ₁	1	5605.93	5605.93	53.78***
	C ₂	1	1728.72	1728.72	16.59***
	Within Samples	297	30968.00	104.27	
	Total	299	48303.00		
Santa Fe	Between Samples	4	4526.00	1131.50	12.37***
	C ₃	1	3242.06	3242.06	35.45***
	C ₄	1	77.06	77.06	0.84
	C ₅	1	620.81	620.81	6.79*
	C ₆	1	586.07	586.07	6.41*
	Within Samples	140	12805.00	91.46	
	Total	144	17331.00		
Main	Between Samples	6	24933.00	4155.50	25.82***
	C ₇	1	83.96	83.96	0.52
	C ₈	1	6148.36	6148.36	38.20***
	C ₉	1	8213.40	8213.40	51.03***
	C ₁₀	1	947.38	947.38	5.89*
	C ₁₁	1	6313.09	6313.09	39.22***
	C ₁₂	1	3226.67	3226.67	20.05***
	Within Samples	203	32674.00	160.96	
	Total	209	57607.00		

TABLE 25
ANALYSIS OF VARIANCE FOR WIDTH OF HEAD

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	6798.00	3399.00	10.38***
	C ₁	1	4368.60	4368.60	13.34***
	C ₂	1	2429.04	2429.04	7.42**
	Within Samples	297	97245.00	327.42	
	Total	299	104043.00		
Santa Fe	Between Samples	4	12404.00	3101.00	17.29***
	C ₃	1	8113.25	8113.25	45.24***
	C ₄	1	160.06	160.06	0.89
	C ₅	1	40.01	40.01	0.22
	C ₆	1	4090.68	4090.68	22.81***
	Within Samples	140	24970.00	179.35	
	Total	144	37374.00		
Main	Between Samples	6	72251.00	12041.83	29.05***
	C ₇	1	6525.00	6525.00	15.74***
	C ₈	1	22736.27	22736.27	54.85***
	C ₉	1	19983.75	19983.75	48.21***
	C ₁₀	1	983.40	983.40	2.37
	C ₁₁	1	12037.69	12037.69	29.04***
	C ₁₂	1	9984.60	9984.60	24.09***
	Within Samples	203	84146.00	414.51	
	Total	209	156397.00		

TABLE 26
ANALYSIS OF VARIANCE FOR HEAD LENGTH/SNOUT LENGTH

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	18340.00	9170.00	9.92***
	C ₁	1	11943.88	11943.88	12.94***
	C ₂	1	6395.80	6395.80	6.93**
	Within Samples	297	274223.00	923.34	
	Total	299	292563.00		
Santa Fe	Between Samples	4	13779.00	3446.75	5.28***
	C ₃	1	3026.24	3026.24	4.64*
	C ₄	1	88.81	88.81	0.14
	C ₅	1	10480.81	10480.81	16.05***
	C ₆	1	183.14	183.14	0.28
	Within Samples	140	91403.00	652.86	
	Total	144	105182.00		
Main	Between Samples	6	87187.00	14531.17	13.39***
	C ₇	1	386.57	386.57	0.36
	C ₈	1	19158.05	19158.05	17.65***
	C ₉	1	1135.35	1135.35	1.05
	C ₁₀	1	580.14	580.14	0.53
	C ₁₁	1	60904.00	60904.00	56.11***
	C ₁₂	1	5023.35	5023.35	4.63*
	Within Samples	203	220339.00	1085.41	
	Total	209	307526.00		

TABLE 27

ANALYSIS OF VARIANCE FOR HEAD LENGTH / EYE WIDTH

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	56509.00	28254.50	53.4***
	C ₁	1	39414.62	39414.62	74.50***
	C ₂	1	17094.00	17094.00	32.31***
	Within Samples	297	157128.00	529.05	
	Total	299	213637.00		
Santa Fe	Between Samples	4	36163.00	9040.75	7.71***
	C ₃	1	10850.74	10850.74	9.26**
	C ₄	1	944.06	944.06	0.80
	C ₅	1	13801.66	13801.66	11.77***
	C ₆	1	10566.54	10566.54	9.01**
	Within Samples	140	164122.00	1172.30	
	Total	144	200285.00		
Main	Between Samples	6	45967.00	7577.83	10.50***
	C ₇	1	23283.96	23283.96	32.27***
	C ₈	1	11456.09	11456.09	15.88***
	C ₉	1	7437.07	7437.07	10.31**
	C ₁₀	1	1152.04	1152.04	1.60
	C ₁₁	1	700.14	700.14	0.97
	C ₁₂	1	1938.02	1938.02	2.69
	Within Samples	203	146481.00	721.58	
	Total	209	192448.00		

TABLE 28
ANALYSIS OF VARIANCE FOR HEAD LENGTH / HEAD WIDTH

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	4829.00	2414.50	21.32***
	C ₁	1	504.17	504.17	4.45*
	C ₂	1	4324.50	4324.50	38.19***
	Within Samples	297	33631.00	113.24	
	Total	299	38460.00		
Santa Fe	Between Samples	4	1847.00	461.75	9.02***
	C ₃	1	152.19	152.19	2.97
	C ₄	1	390.15	390.15	7.62**
	C ₅	1	299.26	299.26	5.85*
	C ₆	1	1005.40	1005.40	19.64***
	Within Samples	140	7166.00	51.18	
	Total	144	9013.00		
Main	Between Samples	6	6137.00	1022.83	9.03***
	C ₇	1	2016.07	2016.07	18.32***
	C ₈	1	0.36	0.36	0.003
	C ₉	1	176.73	176.73	1.61
	C ₁₀	1	1960.00	1960.00	17.81***
	C ₁₁	1	96.80	96.80	0.88
	C ₁₂	1	194.40	194.40	1.77
	Within Samples	203	22335.00	110.02	
	Total	209	28472.00		

TABLE 29

ANALYSIS OF VARIANCE FOR PEDUNCLE DEPTH / DEPTH OF BODY

Series	Source of Variation	D. F.	Sum of Squares	Mean Square	Variance Ratio
Special	Between Samples	2	52784.00	26392.00	30.94***
	C ₁	1	11828.16	11828.16	13.87***
	C ₂	1	40955.22	40945.22	48.01***
	Within Samples	297	253348.00	853.02	
	Total	299	306132.00		
Santa Fe	Between Samples	4	47873.00	11968.25	23.26***
	C ₃	1	21294.72	21294.72	41.38***
	C ₄	1	8096.81	8096.81	15.74***
	C ₅	1	16401.06	16401.06	31.87***
	C ₆	1	2080.41	2080.41	4.04*
	Within Samples	140	72038.00	514.55	
	Total	144	119911.00		
Main	Between Samples	6	47934.00	7989.00	14.70***
	C ₇	1	7409.14	7409.14	13.63***
	C ₈	1	21868.09	21868.09	40.22***
	C ₉	1	7260.00	7260.00	13.35***
	C ₁₀	1	520.80	520.80	0.96
	C ₁₁	1	10155.02	10155.02	18.68***
	C ₁₂	1	721.06	721.06	1.33
	Within Samples	203	110360.00	553.64	
	Total	209	158224.00		

BIOGRAPHICAL SKETCH

John F. Howell was born in Cleveland, Mississippi, on August 2, 1926.

He attended Milligan College, Tennessee, as an officer candidate in the Navy V-12 Unit. Upon discharge from the Navy, he completed the requirements for a Bachelor of Science degree at Delta State College, Cleveland, Mississippi, and received that degree in June, 1948.

He began graduate work in zoology at Louisiana State University in September, 1948, and was awarded the degree of Master of Science in August, 1950. He was elected to the Honor Society of Phi Kappa Phi by the Louisiana State University Chapter in March, 1951.

After being employed by the U. S. Fish and Wildlife Service from August, 1950, until March, 1954, he entered the University of Florida in September, 1954, as a National Science Foundation Fellow. During his tenure at the University of Florida, he has held fellowships awarded by the College of Arts and Sciences, the Graduate Council, and the National Science Foundation. He has also been employed as research assistant, graduate assistant, teaching assistant, and interim instructor by the University of Florida. During the 1957-1958 academic year, he served as Head of the Science Department of Santa Fe High School,

Alachua, Florida, and is currently employed as Assistant Professor of Biology, Austin College, Sherman, Texas.

He has been elected to membership in Phi Sigma Biological Society and associate membership in the Society of Sigma Xi at the University of Florida, where he is now a candidate for the degree of Doctor of Philosophy to be awarded at the June, 1960, commencement.

This dissertation was prepared under the direction of the chairman of the candidate's supervisory committee and has been approved by all members of the committee. It was submitted to the Dean of the College of Arts and Sciences and to the Graduate Council and was approved as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

June 6, 1960



Dean, College of Arts and Sciences

Dean, Graduate School

SUPERVISORY COMMITTEE:

Archie Carr
Chairman

John Kilby
E. Ruffin Jones
Henry M. Wallbourn
A. T. Wallace